Teaching English as a Second Language Quarterly (TESLQ)
(Formerly Journal of Teaching Language Skills)
42(2), Spring 2023, pp. 111-148
HTTPS://DOI.ORG/10.22099/TESL.2023.46955.3177
HTTPS://DORL. NET/DOR/20.1001.1.20088191.2023.42.2.5.5

Online ISSN: 2717-1604
Print ISSN: 2008-8191
Research Paper

Computer-Generated vs. Direct Written Corrective Feedback and Iranian EFL Students' Syntactic Accuracy and Complexity

Reza Bagheri Nevisi [*] **Nikoo Arab** [**]

Abstract

To date, with the everyday growth of technology and the increase of online classes, busy modern teachers seek to lighten their burdens and boost their learners' autonomy. Thus, this multi-method action research aimed to probe the differences between receiving computer-generated and direct corrective feedback (CF) on the syntactic accuracy and complexity of female adolescent Iranian EFL learners' writing. Two intact classes took part in the study; one group (Group C) received CF from an AWE (Automated Writing Evaluation) tool, called Ginger software, and the other group (Group T) received CF from their teacher. Subject-verb agreement and verb form errors were considered to measure the syntactic accuracy while the average sentence length and verb form variation were regarded as the syntactic complexity. Moreover, via a questionnaire, the students' perceptions of the kind of CF they had received were taken into account to see which one was more effective. Findings obtained from MANOVA revealed that both groups made significant improvements regarding syntactic accuracy. Group C produced more complex outputs after being exposed to the treatment while Group T didn't make such progress. The results obtained from the questionnaire also indicated that both groups found the CF satisfactory. However, Group C held more positive attitudes. Not only was the Ginger application effective in decreasing the number of students' writing errors but also it motivated them to be better writers and write more. The study further points to the likely merits of technology-enhanced feedback in improving EFL learners' writing ability in general, and their written syntactic accuracy/complexity in particular.

Keywords: Computer-generated Feedback, Direct Written Corrective Feedback, Syntactic Accuracy, Syntactic Complexity

[*] Received: 03/03/2023 Accepted: 31/05/2023

* Assoicate Professor, Department of English Language and Literature, Faculty of Humanities, University of Qom, Qom, Iran, re.baghery@gmail.com, Corresponding Author
** Instructor, Department of English Language and Literature, Faculty of Humanities, University of Qom, Qom, Iran, arabn.1440@gmail.com

How to cite this article:
Bagheri Nevisi, R., & Arab, N. (2023). Computer-Generated vs. Direct Written Corrective Feedback and Iranian EFL Students' Syntactic Accuracy and Complexity: A Multi-Method Probe into their Effects and Students' Perceptions. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills), 42*(2),111-148. https://doi.org/10.22099/tesl.2023.46955.3177

To date, with the everyday growth of the internet and technology, especially during the pandemic of Covid-19 and increasing online classes, busy modern teachers seek new golden opportunities to lighten their loads in addition to promoting their learners' autonomy. One of the teachers' demanding burdens is providing corrective feedback on students' written production. Both teachers and students find getting feedback on writing necessary (Ferris & Hedgcock, 2005). Being exposed to a language per se does not seem to be enough. That is when corrective feedback plays a crucial role as a learner must become aware of the gap between their own production and the second language rules. Furthermore, obtaining feedback is actually motivating and encouraging (Huxham, 2007).

While it seems necessary to feedback on students' writing, teachers run into some problems. For instance, sometimes teachers are not able to realize what causes the interference from students' first language (Milton, 2006). Moreover, they might even provide subjective feedback (Kaivanpanah et al., 2020). Besides, it might be problematic to offer accurate feedback on students' writing. Novice teachers, particularly, might feel anxious when they are not sure where to begin or how to comment on students' compositions (Ferris & Hedgcock, 2005). On some occasions, learners might misunderstand or even disagree with the feedback provided by their teacher (Ferris, 1995). As a consequence, they might tend to ignore or remove their teacher's feedback (Hyland, 1998). Definitely, if students self-study more and try to discover the errors in their own writing, teachers do not have to spend much time correcting their students' writing (Mohammad Hoesseinpur, (2015); Mohammad Hosseinpur et al., (2018)

There is an alternative to the traditional way of teacher feedback that is computer-generated corrective feedback. However, some teachers may not be familiar with the exact functions of various tools that can ease teaching. Or rather, they tend to stick to the traditional ways of giving feedback because they might not prefer having a process of trial and error. Moreover, for some learners, receiving computer-generated corrective feedback may equal a lack of human interaction (Dikli, 2010). By googling the phrase "best grammar checkers," it can be seen that most websites recommend Ginger software as one of their top AWE programs. However, to the best knowledge of the researchers, there are only a few research studies that have primarily focused on the possible effects or drawbacks of this software. In addition, there are not many studies on the performance of Automated Writing Evaluation (AWE) tools on Iranians, especially adolescents. As a matter of fact, in the context of Iran, to improve EFL students' writing, teacher-generated feedback is frequently provided (Salavatizadeh & Tahriri, 2020). Therefore, this investigation attempted to find out if working with Ginger software could be more useful for this age range compared to getting traditional direct corrective feedback (CF).

Literature Review

Corrective Feedback

Receiving corrective feedback (CF) is very important as it helps students learn through assessment. it is also encouraging and can increase their self-confidence (Alavi,

et al., 2019). Being offered by CF signals students that there is someone who paid close attention to their utterances and sent customized information (Brookhart, 2008). Besides, CF allows instruction to meet the needs of each learner (Ferris & Hedgcock, 2005). As a matter of fact, responding to students' writings is a complex task since teachers themselves have to determine what to consider as an error, how to highlight it, and how offering CF matches other instructional decisions (Ferris & Hedgcock, 2005). These can be added to Zamel's (1985) claim that unfortunately, teachers do not have a guideline when submitting feedback on learners' writing. By browsing the literature, one will realize that the term corrective feedback has adopted different terminology such as Negative Feedback (Aljaafreh & Lantolf, 1994), Written Corrective Feedback (WCF), Written Error Correction or Grammar Correction (GC) (Ferris, 2012; Truscott, 1996).

As stated by Russell and Spada (2006), corrective feedback refers to the feedback students get, which is the sign of an existing error according to language form. Ellis and Shintani (2014) also mentioned that CF is a kind of reaction to students' inaccurate production. In that sense, it might provide negative evidence by highlighting the students' written errors. Nonetheless, it is worth mentioning that correction should not be interpreted as rectifying everything. It should be done to pave the way for students to learn how to get better in putting into words (Edge, 1989). Furthermore, of course, not all errors might be avoided after receiving corrective feedback. According to Ferris (1999), some errors including subject-verb agreement, run-ons, comma splices, missing articles, and verb form errors are categorized as treatable errors. These errors actually occur in a "patterned, rule-governed way" (p. 6). Ferris (1999) argued that the power of feedback might be based on three factors. These factors include teacher, student, and feedback mode, for instance, the instructor's ability to suggest accurate feedback, learners' L1 background or their level of language proficiency, and offering comprehensive, selective, direct, or indirect feedback. Van Beuningen et al. (2012) assumed that CF improves writing accuracy in addition to its long-lasting learning effect. Furthermore, Ferris (2012) stated that CF helps students rectify and edit their writing which leads to improved accuracy over time. Errors and mistakes both need to be corrected for the sake of interlanguage development (Ellis & Shintani, 2014). Moreover, not only students but also teachers appreciate feedback because having written accuracy seems vital in the real world. Actually, instructors must constantly decide which type of corrective feedback works best for their students. There are different types of CF. Sometimes, they only focus on the content of students' writing i.e., the ideas or the messages of the text and even how these ideas were organized. Teachers can also give feedback on form i.e., vocabulary or grammar rules. Feedback can be offered on mechanics either. In that case, it concentrates on the errors in punctuation, spelling, and handwriting. Moreover, CF can be categorized in terms of its focus. If teachers offer unfocused, extensive or comprehensive CF, it means the CF targets various linguistic forms, while focused, intensive, or selective CF signals one or few linguistic forms (Nassaji & Kartchava, 2021).

Another category of feedback is direct and indirect feedback. When teachers write the correct form of the error on students' writing output, direct feedback is actually provided (Ferris, 2003). Indirect feedback, per contra, is offered when the teacher somehow indicates the errors in the learners' utterances (Ferris, 2003). Teachers might provide indirect CF by underlining

the inaccurate productions, writing special codes to specify the type of error, using colors to indicate codes, or writing some comments in the margins of the text (Nassaji & Kartchava, 2021). Actually, this type of feedback encourages students to engage and self-correct their own writings. Hence, it is not suitable for lower proficiency students since they seem unable to identify the correct form of their errors (Ferris & Hedgcock, 2005). According to Ferris (2006), learners prefer to get direct feedback because they think it is more effective than receiving indirect feedback. It is to some extent because they have to simply copy the correct form provided by the teacher. Chandler (2003) also suggested that direct CF is preferred since it is the quickest and simplest way for both teachers and students.

Syntactic Accuracy and Complexity

Truscott (1996), Polio et al. (1998), and Fazio (2001) all agreed that grammar correction and offering corrective feedback might not only be ineffective but also disheartening for language learners. According to Truscott (1996), teachers have to spend a lot of time on error correction. This wastes their energy and distracts them from more valuable issues in class. Many instructors, however, deem that students need to receive corrections from the teacher so that they can improve their writing accuracy i.e., they write a text that complies with L2 language rules (Yuan & Ellis, 2003). Actually, offering written CF can help students strengthen their writing ability. Taking the literature into consideration, defining complexity seems to be a highly controversial issue. Traditionally, when defining Grammatical Complexity (or Syntactic Complexity) people considered a sentence as a complex one based on the number of embedded subordinate clauses it contained (Biber & Gray, 2016). According to Foster and Skehan (1996), complexity indicates students' eagerness to restructure when their language is advancing toward more complex subsystems. However, it is worth noting that there are different types of syntactic complexity. As a matter of fact, they are dissimilar in terms of the syntactic, discourse, and register features they have (Biber & Gray, 2016). For instance, as these scholars believe, instead of considering embedded subordinate clauses, the number of embedded phrasal structures can be taken into account.

Ellis and Barkhuizen (2005) categorized complexity into interactional, propositional, functional, lexical, and grammatical types. Some scholars relate grammatical complexity to the number of clauses per T-unit (e.g., Larsen-Freeman, 2006) while others defined it as the ability to use a specific structure (e.g., Yuan & Ellis, 2003). Moreover, researchers such as Ortega (2003) and Fazilatfar et al. (2014) considered the mean length of sentence (MLS) and the ratio of dependent clauses per clause (DC/C) as the syntactic complexity. Learners might use simple (as opposed to complex) and well-structured language forms to produce L2 output. While Accuracy focuses on language form and is related to a specific interlanguage level of students, Complexity reflects students' keenness to take risks and is related to the progress of their interlanguage system (Foster & Skehan 1996).

Related Studies

Although studies on teacher written feedback and computer-generated feedback abound in the literature, few have delved into the effect of TF and CF on the syntactic complexity and accuracy of students' writings within a single study. Here are some relevant studies that have been conducted on the above-mentioned variables. Chandler (2003) probed the accuracy and complexity of 31 ESL college students. By offering them different types of feedback, he found out that although the accuracy of the experimental group showed improvement after being exposed to the treatment, the complexity was not affected. This study specifically dealt with students' accuracy and complexity in an ESL context. In the same vein, Sheen (2007) investigated 91 ESL learners who had different L1 backgrounds. There were three groups. One of them received direct feedback, the other received direct metalinguistic while the last group received no feedback. It revealed that the writing of both experimental groups was significantly improved. Fathman and Whalley (2012) also studied 72 students with different L1 backgrounds but with the same language proficiency at two colleges. The results of this study indicated that receiving corrective feedback on grammar and content help students develop their writing skill. Moreover, without getting feedback, many students could write more after rewriting their own texts. Van Beuningen, et al. (2012) conducted a study on 268 students to see the short-term and long-term changes between getting direct and indirect CF. To measure accuracy, they calculated the number of grammatical errors divided by the total number of words. To analyze complexity, they measured the number of subordinate clauses divided by the total number of clauses. The findings suggested that both types of feedback were effective in improving students' accuracy. Moreover, receiving CF did not lead to any significant differences between the groups in terms of lexical and syntactic complexity.

Farrokhi and Sattarpour (2012) conducted an experiment on 60 advanced EFL students. The results indicated that the grammatical accuracy of the experimental groups who had received focused and unfocused feedback significantly improved. The results also revealed that offering focused feedback would be much more effective. Fazilatfar et al. (2014) investigated the impact of unfocused written CF on the lexical and syntactic complexity of the writing of 30 male and female advanced students. He suggested that the syntactic and lexical complexity of the experimental group had significantly increased. These two research mainly concentrated on focused and unfocused feedback and this is what distinguishes them from similar studies done in the literature.

Kang and Han (2015) adopted a meta-analytic approach to finding out the possible effects of written CF on the grammatical accuracy of the students' writing. This study also showed that written CF facilitates students' L2 writing performance. Furthermore, a more recent quasi-experimental study done by Valizadeh (2020) on 90 Turkish EFL learners indicated that either giving direct CF or metalinguistic explanations to students could have positive effects on their written performance, particularly their syntactic accuracy. A very recent experiment was done on 60 Iranian elementary EFL learners (Valizadeh & Soltanpour, 2021). One of the experimental groups was offered focused

direct feedback while the other group received no kind of CF during a three-week treatment. However, the post-test results of those groups were not significantly different in terms of syntactic complexity.

The following studies specifically addressed computer-generated feedback on students' writings via various tools and software, students' perceptions of such feedback, and their possible merits and demerits. Students in Chen and Cheng's (2006) investigation reported a negative perception of the computer-generated feedback they had received from an AWE tool, namely My Access. These 68 students believed that this program is unable to offer personalized CF and is not a capable judge.

A quasi-experimental study on 12 adult ESL students who worked with My Access, suggested that feedback provided by teachers or computers is unalike in nature (Dikli 2010). While the former is much briefer and to the point, the latter is redundant and unneeded. Therefore, it can be confusing for lower-level students. Moreover, Lai's (2010) study on 22 Taiwanese EFL students who worked on another software, called MY Access, indicated no significant improvement in the learners' writing. In fact, after submitting their draft to this grammar-checker, their peers revised their utterances in the classroom. These participants thought highly of being evaluated by their peers. In addition, 45 Taiwanese college students who obtained computer-generated feedback through the My Access program were also investigated (Fang, 2010). Boosting the students' motivation and autonomy led them think highly of this program and like to have it as a support in their future writing classes. In their large-scale three-year research on more than 8 thousand students, Grimes and Warschauer (2010) found out that using AWE tools eased classroom management, avoided wasting class time, and boosted the learners' autonomy and motivation to write and revise more. Results from Dikli and Bleyle (2014) who had focused on Criterion software as well, suggested that all 14 ESL college students had positive perceptions. Furthermore, Cavaleri and Dianati's (2016) study indicated that Australian higher education students have positive perceptions towards Grammarly since it was easy to use and it helped them improve their writing.

A mixed-method study was done to investigate the possible impacts of direct and metalinguistic electronic feedback on 29 Iranian EFL learners' writing accuracy (Saadi & Saadat, 2015). One group received direct electronic corrective feedback from Ginger software, while the other group was offered metalinguistic electronic corrective feedback provided by another software, called Markin4. The findings of this study indicated that receiving both types of electronic feedback would improve learners' writing accuracy. The students found out that being corrected by their teachers was not pleasant because of their silly mistakes. Additionally, receiving computer-generated corrective feedback helped them have a student-centered class and this boosted their autonomy but even so, some of them wished to have it as a supplementary to their teacher's feedback.

Li et al. (2015) probed 70 ESL university students to find out the students' and instructors' views on the corrective feedback generated by the Criterion software and its effect on the students' accuracy. The results revealed that the instructors were not happy with the quality of the feedback but agreed on the fact that it helped out the learners in

grammar and mechanics. Students' view was generally positive, and receiving this type of feedback assisted them in order to improve their writing accuracy.

There are also various studies examining users' perceptions of computer-generated CF. For instance, Qassemzadeh and Soleimani (2016) conducted research on 70 intermediate male and female EFL learners to discover the effect of feedback provision by teachers and Grammarly software on learning passive structures. They concluded that receiving Grammarly's feedback helped more than traditional teachers' feedback in terms of long-term retention of passive structure.

Huang and Renandya (2018) also analyzed the effect of automated feedback provided by Pigai software on 35 Chinese EFL university students who were all at a low level of language proficiency. A control group of 32 students was selected randomly as well. Although the writing quality of the experimental group did not improve significantly, the quantitative phase of the study showed that they have a very positive attitude toward using this program. This was in line with Lu's (2019) idea that with Pigai, learners prompt to practice writing more and more.

The results of a four-week investigation (Hajebi, 2018) conducted on 74 Iranian EFL university students indicated that getting computer-generated CF by Ginger software significantly improved students' writing performance. Consequently, the researcher believed that guiding students in the class could pave the way for them to comfortably work with computers in order to learn at home easier.

O'Neill and Russell (2018) also investigated 96 university students and compared the impacts of Grammarly's computer-generated feedback with non-automated grammar feedback they had actually received from their instructor, using Track Changes of Microsoft Word. These students showed more satisfaction with Grammarly. Yet, since the program was not accurate enough to guarantee independent use, they suggested using it in combination with the instructor's advice.

Kaivanpanah et al. (2020) attempted to discover the difference between receiving computer-generated feedback and teacher feedback on writing quality. They chose two grammar checkers, namely Grammarly and Write & Improve. The findings of their investigation revealed that in general, teacher feedback had been more helpful in improving the students' writing quality.

Through mixed-method research, Salavatizadeh and Tahriri (2020) investigated the impact of blending computer-generated feedback and teacher feedback on 30 EFL learners' writing. The control group only received teacher feedback. Another purpose of their research was to see the learners' attitudes about receiving both types of feedback. The results of the experimental group were generally better than the results of the control group. Furthermore, the learners showed a favorable attitude towards being offered feedback by both automated and teacher feedback.

In a very recent study, Wang and Han (2022) compared the results of 70 Chinese EFL students who had received either their teacher's or automated corrective feedback. They found out that the learners who obtained automated feedback outperformed their counterparts in the post-test. Nonetheless, teacher feedback had more psychologically-

positive impacts. To sum up, it is not clear whether receiving computer-generated feedback truly satisfies all learners or not. Research done in this area has its own limitations and some gaps can be observed in the existing literature. First of all, most of the research conducted on computer-generated feedback only includes ESL learners. Second, none of these experiments have concentrated on adolescents. More importantly, none of them have probed students' writing complexity when using a grammar checker. Therefore, to fill the above-mentioned gaps, the researchers formulated these questions:

1. Is there a significant difference between receiving computer-generated and direct corrective feedback on the syntactic accuracy of EFL learners' writing?
2. Is there a significant difference between receiving computer-generated and direct corrective feedback on the syntactic complexity of EFL learners' writing?
3. What are the EFL learners' perceptions of computer-generated and direct corrective feedback they receive?

Method

Participants

The participants of the study were 26 female EFL students of a private junior high school in Tehran. Therefore, all the participants were Iranian and their age range was 13-14. Two intact classes were chosen as the experimental groups of the study. The first experimental group used an AWE software, called Ginger (n = 14). The second one received their teacher's direct CF (n = 12). In the present research the first group that received computer-generated feedback is called Group C and the other group that received teacher's feedback is called Group T. Moreover, another 25 students from two different classes responded to the questionnaire to help pilot the questionnaire and confirm its reliability. Students of this junior high school attend a general English course and it is actually a part of their school curriculum. Both classes were studying the same units of the book, World English 1. According to CEFR (Common European Framework of Reference for Languages) levels, this book is considered as A2+ or pre-intermediate. Nevertheless, in order to ensure the homogeneity of these groups, the researchers administered the KET Reading and Writing Sample Test which was featured for this proficiency level and the age range before the beginning of the treatment.

Instruments

KET Reading and Writing Sample Test

Key English Test (KET) is the basic level of the General English exam in the Cambridge English range. As a matter of fact, since the level of this test is considered A2+ (or pre-intermediate) according to CEFR, its score indicates whether one can communicate in basic English in day-to-day life or not. At this level, a learner should be able to use English to contact native or non-native speakers and understand newspapers, leaflets, posters, street signs, city guides, etc. (University of Cambridge, 2007). For the present research, KET Reading and Writing sample test was chosen to measure the participants' English level of proficiency. In addition to the fact that this test is suitable

to the age range of the participants, based on its features, it seemed appropriate to be used to measure the homogeneity of the groups of participants. The exam administered for the students of this research contained 7 parts with 32 questions in total within a time limit of one hour. In other words, it has 30 reading questions together with two writing parts. It generally includes cloze tests, choosing the right response to a dialogue, understanding the main idea and details of a text, and writing short messages. For each correct answer in parts 1 to 5, students get one mark and a maximum of 15 marks each for the last two parts (60 points in total).

Picture Stories

In total, six picture stories were used during the present research; two for the pre-and the post-test, and four picture stories were given to the students as their writing assignments during the treatment. These picture stories were collected from the Young Learners English (YLE) tests or YLE activity books. Although the University of Cambridge designs YLE tests for four- to twelve-year-old children, the pictures seemed suitable and interesting for the participants of this study. They are available in appendices B, C, D, E, F, and G.

AWE Program: Ginger

Ginger Software (gingersoftware.com) is an AWE program that was first founded in 2010. This AWE software enjoys an Artificial Intelligence (AI) system that is able to check the accuracy of a text by reviewing the sentence in addition to suggesting some replacements for the errors immediately. At first, it uses the color red to highlight and underline the errors found in an essay. Then, with a single click, all the errors can be corrected. The Ginger program contains a Sentence Rephraser as well. This feature helps users modify a sentence with inspiring phrases and even idioms. It is more beneficial when people would like to write clearer and more engaging texts. Another positive feature of this program is providing AI-based synonyms. By double-clicking a word, one can see a list of various synonyms of that word. Therefore, there is probably no need to look up a dictionary while reading a text on a website. This software offers two premium plans called Ginger for Education and Ginger for Business in addition to a free plan which consists of basic writing suggestions i.e., providing synonyms and antonyms, correcting the grammar, spelling, and punctuation explicitly which definitely improves everyday communication. Ginger for Education explains the errors found in an essay. This feature actually provides a learning opportunity. A manager might like to use Ginger for Business for his/her one hundred thousand team members. AI-powered translation is also provided by this program which surprisingly supports 40 languages. Thus, everyone in a company can make almost error-free texts. Furthermore, employees, customers, or students with different language backgrounds can use AI-powered translation to understand the texts. The participants were all adolescents so in addition to availability, the comfortability of the software had been taken into account. Ginger software is available as web-based, a

desktop app, a mobile app, a Chrome extension, and a Microsoft Word add-in. Its user-friendly app is also easy to install from App Store and Google Play. Moreover, its appearance is either simple or modern. At first, only the Chrome extension was introduced to the participants because in that case, they were able to type easily on their Learning Management System (LMS), called Moodle, while Ginger automatically suggested to them feedback in the blink of an eye. Besides, since one of the purposes of the study is to check subject-verb agreement and verb form errors, using the free version was appropriate.

Questionnaire

To answer the third question of this research and explore the participants' perceptions of using Ginger and receiving direct CF, a particular questionnaire was given to each group of students after the treatment. All the items in the questionnaire were inspired by previous research on computer-generated feedback (e.g., Huang & Renandya, 2018). Students had to choose one among *Strongly Agree*, *Agree*, *Neutral*, *Disagree*, and *Strongly Disagree*. Both questionnaires are available at Appendices H and I. The reliability of the questionnaires was confirmed by piloting it on 25 students. Moreover, their validity was confirmed by two university instructors. These questionnaires were supposed to delve into five items: The students' general satisfaction with the CF, the effectiveness of the CF in eliminating their errors, the comprehensibility of the CF itself, the sense of encouragement caused by CF, and their desire to receive that particular CF again in the future.

Data Collection Procedure

To begin with, KET Reading and Writing Sample Test was administrated for approximately an hour to see if both groups were homogeneous enough. Next, as a writing pre-test, the learners were given a picture story. Then, the teacher asked them to write 40-50 words about it. In order not to manipulate the results, the students' oral questions were not answered. Instead, the teacher asked them to write based on their own prior knowledge and use any words and structures they know. Only the subject-verb agreement errors, verb form errors, sentence length, and verb form varieties were spotted by the researchers. For instance, in the sentence, *a puppy run over the car*, there is a subject-verb disagreement. Furthermore, in the sentence, *now the car clean*, the verb is missing, and in *they ate fruit juice*, the wrong verb was used. Both these errors were considered verb form errors. Moreover, to measure verb form varieties, different tenses of verbs were considered. For example, simple present, present perfect, or even modal verbs were each considered as a variety. During the treatment, the two groups wrote or typed a paragraph based on different picture stories once a week (40-50 words). Group C who received Ginger's feedback had two options to submit their weekly assignments. They could either type and send their task to the teacher on their LMS, called Moodle or give it to their teacher in person. The second group, group T could bring their writing assignments to their classroom and give them to the teacher. To indicate the possible improvements in

writing accuracy, group C did not receive any feedback, but from the AWE software, i.e., Ginger, while the teacher of the other group offered them direct corrective feedback on their errors. Seeing that the Ginger program provides comprehensive CF as opposed to selective feedback, the teacher of group T gave comprehensive CF as well. In other words, that teacher did not only offer feedback on verb form or subject-verb agreement errors. So as to recognize the possible improvements in the writing complexity, group C used the Ginger Sentence Rephraser to see some other options for each of their phrases and sentences. Furthermore, the teacher of group T suggested some alternatives to the students' phrases and sentences. As an example, when a student wrote, *Sarah's family likes to go to the park but their car is very dirty*, the teacher suggested, *Sarah's family are going to go to the park but their car looks very dirty*.

Table 1
Research Procedure Timetable

Session	Procedure
1st	The KET was administrated.
2nd	The pre-test was administrated.
3rd	1st writing homework was given.
4th	No assignments were given.
5th	2nd writing homework was given.
6th	No assignments were given.
7th	3rd writing homework was given.
8th	The questionnaire was piloted.
9th	4th writing homework was given.
10th	No assignments were given.
11th	The post-test was administrated. The questionnaire was carried out as well.

After four weeks of receiving computer-generated or teachers' corrective feedback, each group had a writing post-test to be compared with the pre-test in terms of syntactic accuracy and syntactic complexity. Again, 40-50 words had to be written about a specific picture story. Appendix G demonstrates the post-test. Afterward, another 25 students of this school participated in the piloting phase of the study in which the questionnaire was piloted to ensure the reliability. At last, all the collected data from groups C and T were triangulated with a questionnaire in order to find out the learners' perceptions of the type of feedback they had received.

Data Analysis

The data collected from the pre-test and the post-test of the participants were computed in the latest version of SPSS Statistics, which is 26. Using the Shapiro-Wilk Test, the normality of groups T and C was proved. Next, Levene's Test was used to see

whether the groups were homogenous or not. Then, Multivariate ANOVA (MANOVA) was used to discover whether there were any significant differences in the performance of the two experimental groups after being exposed to the treatment. Moreover, in order to find the answer to the third question, the responses to the questionnaires were analyzed by measuring the mean of responses.

Results

Homogeneity of the Groups

Table 2 presents the descriptive statistics of groups T and C. As seen in this Table, the mean and the standard deviation of these two groups are very similar.

Table 2
Group Statistics in KET Test

	N	Mean	Std. Deviation	Std. Error Mean
Group T	12	54.08	1.083	.312
Group C	14	54.14	1.167	.311

Prior to the homogeneity check, since there was a small sample size, Shapiro-Wilk Test measured the normality of the two groups, using the students' scores in KET Reading and Writing Sample Test. As shown in Table 3, Shapiro-Wilk Test did not indicate any evidence of non-normality since the p-value of both groups is greater than 0.05.

Table 3
Test of Normality in KET Test

	Shapiro-Wilk		
	Statistic	**df**	**Sig.**
Group T	.939	12	.487
Group C	.936	14	.370

Next, Levene's Test was run to make sure about the homogeneity of groups T and C. The p-value in this test was 0.671 and obviously greater than 0.05. This result revealed that these groups were not significantly different from one other (Table 4).

Table 4
Levene's Test for Equality of Variances in KET Test

F	**Sig.**	**T**	**df**	**Sig. (2- taile-d)**	**Mean Differe-nce**	**Std. Error Differe-nce**	**95% Confid-ence Interval of the Differe-nce**
							Lower
.185	.671	-.134	24	.895	-.05952	.44443	-.97679

Syntactic Accuracy and Complexity of the Groups

Table 5 indicates the descriptive statistics of the pre-and post-test of groups T and C. The four dependent variables of the study are listed in the first column of the table. The mean of the performance of each group in the pre-and post-test was measured. Both groups could successfully decrease their errors in accuracy after being exposed to the treatment. Moreover, group C's results showed that they could produce more complex output due to the increase in their sentence length mean and verb form variation. However, this was not true for group T.

Table 5
Descriptive Statistics of the Pre- and Posttest

Measure	Groups	Time	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Subject Verb Agreement Errors	T	Pre-Test	2.000	.195	1.598	2.402
		Post-Test	.417	.135	.137	.696
	C	Pre-Test	1.929	.180	1.556	2.301
		Post-Test	.214	.125	-.044	.473
Verb Form Errors	T	Pre-Test	2.583	.146	2.282	2.884
		Post-Test	1.917	.164	1.578	2.256
	C	Pre-Test	2.643	.135	2.364	2.922
		Post-Test	.286	.152	-.028	.600
Sentence Length Mean	T	Pre-Test	11.771	.290	11.172	12.370
		Post-Test	11.813	.200	11.399	12.226
	C	Pre-Test	11.375	.269	10.820	11.930
		Post-Test	13.179	.186	12.796	13.562
Verb Form Variation	T	Pre-Test	10.500	.436	9.600	11.400
		Post-Test	10.750	.332	10.065	11.435
	C	Pre-Test	10.143	.404	9.310	10.976
		Post-Test	13.500	.307	12.866	14.134

Furthermore, MANOVA was run to answer the first two questions of this research (Table 6). It should be noted that the between-subjects factor was the groups, i.e., groups T and C. According to the results of MANOVA (p-value = .001), the probability of obtaining these findings by chance was low. In other words, there were significant differences between groups T and C on accuracy and complexity, and for sure these groups were not the same.

Table 6
Multivariate Tests of the Syntactic Accuracy and Complexity

Effect			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Bet...	Intercept	Pillai's Trace	.998	2738.357	4	21	.000	.998
		Wilks' Lambda	.002	2738.357	4	21	.000	.998

Effect			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Groups		Hotelling's Trace	521.592	2738.357	4	21	.000	.998
		Roy's Largest Root	521.592	2738.357	4	21	.000	.998
		Pillai's Trace	.558	6.638	4	21	.001	.558
		Wilks' Lambda	.442	6.638	4	21	.001	.558
		Hotelling's Trace	1.264	6.638	4	21	.001	.558
		Roy's Largest Root	1.264	6.638	4	21	.001	.558

Questionnaires Results

At first, the reliability of the questionnaires had to be ensured. Therefore. Cronbach's alpha was estimated for the questionnaire piloted with the participation of 25 students in the same school. Five things were supposed to be surveyed from the participants of the study: Students' satisfaction with the CF, the effectiveness of the CF in decreasing their errors, the clarity of the CF itself, the increasing motivational level caused by CF, and their desire to receive that particular CF again in the future. As shown in Table 7, Cronbach's alpha for all five elements is greater than 0.7 and consequently, it is assumed that all items in this questionnaire are totally reliable.

Table 7
Piloted Questionnaire – Reliability

	Cronbach's Alpha	N of Items
The Satisfaction with the CF	.942	6
The Effectiveness of the CF in Decreasing the Errors	.871	7
The Clarity of the CF	.890	5
The Increasing Motivational level caused by CF	.867	3
The Desire to Receive CF in the Future	.949	5
Total Items	.923	26

As mentioned before, students had to choose one among *Strongly Agree, Agree, Neutral, Disagree, and Strongly Disagree*, ranging from 0 to 4 respectively. Therefore, after aligning the collected data, if the mean of learners' responses to each category of questions is more than three, it can be concluded that the students agreed on that particular category. Table 8 demonstrates the descriptive statistics of group T. The mean of group T's responses to the first six questions which reveals their general satisfaction with the feedback they got, is 3.09. It suggests that on average, students agreed that the feedback was satisfying. On the other hand, they did not see that feedback was effective enough since the mean of their responses to the next seven questions is 1.67. Additionally, the students assumed that the CF they had received was clear (Mean=3.65) and they would like to be offered again in the future (Mean=3.28) although this particular CF did not encourage them much to write more (Mean=2.19).

Table 8
The Perspectives of Group T - Descriptive Statistics

	N of Items	Min.	Max.	Mean Error	Std.	Std. Deviation
The Satisfaction about the CF	6	2.67	4.00	3.09	.12	.43
The Effectiveness of the CF in Decreasing the Errors	7	1.14	2.43	1.67	.08	.31
The Clarity of the CF	5	3.00	4.00	3.65	.12	.41
The Increasing Motivational level caused by CF	3	1.00	3.00	2.19	.19	.67
The Desire to Receive CF in the Future	5	2.60	3.80	3.28	.12	.43

Furthermore, Table 9 shows Group C's opinion about Ginger's feedback. The mean of all the items is greater than 3.7. Thence, the learners of group C were generally satisfied with the CF (Mean=3.79). It was mainly helpful (Mean=3.87) and understandable (Mean=3.85). In general, it encouraged them to write more (Mean=3.8) and they preferred to be given feedback by Ginger Software (Mean=3.87).

Table 9
The Perspectives of Group C - Descriptive Statistics

	N of Items	Min.	Max.	Mea	Std. Error	Std. Deviation
The Satisfaction about the CF	6	3.50	4	3.79	.04	.17
The Effectiveness of the CF in Decreasing the Errors	7	3.57	4	3.87	.04	.16
The Clarity of the CF	5	3.00	4	3.85	.09	.36
The Increasing Motivational level caused by CF	3	3.00	4	3.80	.08	.31
The Desire to Receive CF in the Future	5	3.40	4	3.87	.05	.20

Discussion

While in general, students appreciate getting CF from their teacher, AWE has always been a subject of debate in the literature. Similar to other technologies, AWE tools seem incapable of making ineffective teaching successful; rather they can be a support for good teachers to make their teaching more fruitful (Grimes & Warschauer, 2010). According to some scholars (e.g., Liao, 2015; Salavatizadeh & Tahriri, 2020), teacher feedback should not be replaced by computer-generated feedback and it is the students' need that determines which type of feedback must be used. Comparing the results of the participants revealed that the number of subject-verb agreement errors of both groups significantly decreased in the post-test. This is also true about their verb form errors in which group C showed even more significant improvement. Consequently, the syntactic

accuracy of all the participants of the present study improved after receiving the treatment.

Qassemzadeh and Soleimani (2016) are other researchers who have reached almost similar results. They declared that although feedback provided by Grammarly software and teachers can both have positive effects on passive structure mastery, Grammarly seemed more helpful. The findings of Saadi and Saadat (2015) who investigated Ginger and another software called Markin4, are in consonance with the present research as well. However, the results do not concur with the findings of Kaivanpanah et al (2020). To begin with, they compared the differences between receiving indirect coded teacher's CF and computer-generated feedback on learners' writing. Although both groups improved when getting CF, one of the groups showed more improvements after receiving the teacher's CF. Seeing that as opposed to the present study, the teacher in that research offered indirect coded CF students had to analyze their teacher's marks. This probably led to their writing quality enhancement.

In the present study, the two groups' results were dissimilar in terms of syntactic complexity. The sentence length mean of group C significantly became greater after the treatment. Meanwhile, group T's sentence length mean did not increase. The same is true for group T's verb form variation. On the contrary, group C's verb form variation was significantly improved in the post-test. According to Hyland (1993), providing students with the opportunity to use AWE tools per se does not make them better writers. Notwithstanding, students of group C had been positively influenced by Ginger software.

As mentioned earlier, there have been no experimental studies that have explored the effects of computer-generated feedback on the syntactic complexity of students' writing. Therefore, only group T's results could be compared to the related studies. Although to some extent the data collection procedure of researchers such as Chandler (2003), Valizadeh and Soltanpour (2021), and Van Beuningen, et al. (2012) is different from this study, the results are aligned. All these researchers found that teachers' feedback does not have an influence on the complexity of their students' writings.

The results of Fazilatfar et al. (2014) who discovered that the syntactic complexity of the experimental group had significantly improved after receiving their teacher's comprehensive feedback, are not in agreement with the findings of this study. One reason is that their data collection procedure completely differed from this study. They defined syntactic complexity as the mean length of the sentence and the dependent clauses per clause. The participants were at an advanced level of English and they were all older than the students in this research. Moreover, the investigation of Fazilatfar et al. (2014) lasted 3 months, 3 sessions a week. Therefore, another probable justification for this disagreement is the longer duration of their research.

As presented in Appendices H and I, five things were surveyed from the students who participated in this research. First of all, they were questioned whether the CF they had received was satisfactory or not. It was discovered that both groups generally found getting CF fulfilling. The results of the questionnaire that was given to group C are in accordance with many investigations such as Cavaleri and Dianati (2016), Dikli and

Bleyle (2014), Fang (2010), Grimes and Warschauer (2010), Huang and Renandya (2018), Li et al. (2015), Lu (2019), and Saadi and Saadat (2015). On the other hand, the finding of this study is not in line with the Taiwanese university students who participated in Chen and Cheng's (2006) experiment. They, in fact, reported their dissatisfaction with the AWE tool they had been using. Since they worked on MyAccess, the students' opinions might totally differ from this study.

The second surveyed thing was about the effectiveness of the CF i.e., whether it was generally helpful in decreasing their errors or not. The groups of the present study did not have the same idea in this regard as group T believed that their teachers' CF had been ineffective. Nonetheless, group C's views are in agreement with the participants of Huang and Renandya (2018).

The next thing was the understandability of the CF. Since learners' language backgrounds differ from one another, CF must be sufficiently clear; in other words, it is necessary for them to comprehend the feedback information as the instructor [or the feedback provider] had it in mind (Brookhart, 2008). All the participants of this study agreed that the CF they had received was comprehensible and clear. This is again in accordance with the study of Huang and Renandya (2018).

Afterward, they were asked whether the CF could have encouraged them to write more. The teacher's CF did not generally motivate group T to become better writers. While group C's students found Ginger's feedback much encouraging. This result is in line with the findings of some researchers such as Fang (2010), Li et al. (2015), and Lu (2019).

In the end, the students answered some other Likert scale questions to see whether they would like to receive that particular CF again in the future. Surprisingly, the results indicated that both groups prefer to obtain CF in the future. Even group T found the teacher's CF ineffective and demotivating. Hence, it can be inferred that all the students in this research find receiving CF as an indivisible part of their education. This conforms to the idea of Ferris and Hedgcock (2005).

For sure, offering direct CF to learners is of extreme importance and each type of CF might have its own strengths and weaknesses. However, it requires time in addition to repetition to be more fruitful for them. In consequence, by giving enough time to students, the chances to show significant improvements could get higher. Thus, teachers can provide opportunities to help their language learners develop in L2 writing by finding students' wants and needs. In conclusion, the nature of the English courses in this Junior High school might not be similar to all other English courses and contexts. The results might vary if the time of the research procedure was longer. However, the results agree with most experiments documented elsewhere.

Conclusion

This research sought to discover if there is a significant difference between receiving computer-generated and written direct corrective feedback on Iranian EFL adolescents' writing syntactic accuracy and syntactic complexity. To do so, only two aspects of

accuracy and two aspects of complexity were analyzed. Nevertheless, the generalizability of findings could be undermined by the fact that the research was conducted in two intact classrooms and due to its small sample size, the study might fall short of being generalizable to all similar contexts.

The conclusion that could be drawn from the findings of this study is that providing feedback is necessary as producing output is one of the teachers' priorities. There is an alternative to the traditional way of offering corrective feedback i.e., guiding students to receive computer-generated feedback and benefit from technology. Not only would it lessen the loads of teachers, but also students can increase their autonomy and flourish at their own pace. Although teachers' CF is a great option to help students write more accurately, their suggestions seem unable to improve students' written complexity. On the other hand, it was proved that Ginger Rephraser can provide learners a chance to learn how to produce more complex utterances. Moreover, the results follow the fact that getting corrective feedback is satisfactory and effective either received from teachers or via Ginger software. However, in comparison to this grammar-checker, students might assume that teachers' CF does not generate that much motivation in them to become better writers.

According to the findings of this research, the following points can be considered as the pedagogical implications. The results of the present research will contribute to stakeholders such as administrators and curriculum developers. Ginger software was regarded as a great support for learners and as an effective tool that aids students to boost their syntactic accuracy and complexity. Therefore, stakeholders may benefit from these findings and decide to make use of grammar-checkers such as Ginger and ask teachers to introduce this software to their learners. Besides, it is clear that it is worth buying the premium plans of Ginger e.g., Ginger for Education, to empower their students. Furthermore, since teacher's CF was found helpful, administrators who do not use this technology must make sure the teachers constantly provide feedback on the students' writings. In addition, because it was found out that Ginger can be a great help to the EFL adolescent participants of the study, AWE tool developers can also profit from this study in that they can make an effort to focus on the strengths of their software and make it even better. There are some implications for language teachers as well. Although receiving feedback from their teacher is necessary and satisfying for EFL adolescents, in order not to waste their time and energy, teachers can introduce Ginger to their students and wait for it to work and then, see the improvements. To summarize, considering all the findings of the study, it is obvious that working with Ginger software would be beneficial for language learners. Not only is it effective in decreasing the number of their writing errors, but also it motivates them to be better writers and write more.

Like all the previous studies, this research had its own limitations. First of all, since this study was done as a natural part of the English course of a junior high school, only intact sampling was practical for the researcher. The second limitation was that the participant's personal variables e.g., the level of openness, initial motivation, anxiety, and autonomy had not been considered. In spite of the teachers' explanations, it is not clear

whether the students paid sufficient attention to the questions of the questionnaire or not. In other words, some of the participants might not have carefully responded to the questionnaire. Since all the assignments had been done at home, the researcher could not make sure whether the students had received any help or used any other resources or not. Last but not the least, this research was carried out in the summer term of their English course, so there was no choice but to investigate the possible short-term effects of computer-generated and teacher's corrective feedback since it was a four-week study.

In the present study, intact sampling was used. It is suggested to use other types of sampling on a larger scale administration and see how the generalizability of findings differs. Future researchers might also like to add a control group to their experiment. This research tried to investigate the possible short-term effects of receiving computer-generated and direct CF on syntactic accuracy and complexity. Therefore, the next research in this area could define syntactic accuracy and complexity in a different way and assess other aspects of them. In addition, the possible long-term effects of these two types of feedback can be investigated. Besides, it is recommended to use other technologies and grammar checkers to find out the possible effects on students' syntactic accuracy, complexity, or even fluency. It was not feasible for the researcher to buy the premium plans of Ginger software. So, empowering students with those plans might reveal another result. Last but not least, using a couple of grammar checkers simultaneously and surveying their advantages and disadvantages might be of interest to some investigators.

References

Al-Ahdal, A. A. M. H. (2020). Using computer software as a tool of error analysis: Giving EFL teachers and learners a much-needed impetus. *International Journal of Innovation, Creativity, and Change*, *12*(2), 418–437.

Alavi, S., Kaivanpanah, S., & Danesh, F. (2019). A comparative study of writing assessment using activity theory-based assessment model (ATBAM) and a traditional approach. *Journal of English Language Teaching and Learning*, *11*(23), 1–25.

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, *78*(4), 465–483. https://doi.org/10.1111/j.1540-4781.1994.tb02064.x

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English linguistic change in writing*. Cambridge University Press.

Brookhart, S. M. (2008). *How to give effective feedback to your students*. Association for Supervision and Curriculum Development.

Cavaleri, M., & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Language and Learning*, *10*(1), 223–236.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, *12*(3), 267–296. https://doi.org/10.1016/s1060-3743(03)00038-9

Chen, C.-F. E., & Cheng, W.-Y. E. (2006). The use of a computer-based writing program: Facilitation or frustration? [Paper presentation], 23rd International Conference on English Teaching and Learning, Republic of China.

Chen, C.-F. E., & Cheng, W.-Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language, Learning and Technology*, *12*(2), 94–112.

Dikli, S. (2010). The nature of automated essay scoring feedback. *CALICO Journal*, *28*(1), 99–134. https://doi.org/10.11139/cj.28.1.99-134

Dikli, S., & Bleyle, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback? *Assessing Writing*, *22*, 1–17. https://doi.org/10.1016/j.asw.2014.03.006

Edge, J. (1989). *Mistakes and corrections*. Longman.

Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford University Press.

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. Routledge.

Fang, Y. (2010). Perceptions of the computer-assisted writing program among EFL college learners. *Educational Technology & Society*, *13*(3), 246–256.

Farrokhi, F., & Sattarpour, S. (2012). The effects of direct written corrective feedback on improvement of grammatical accuracy of high-proficient L2 learners. *World Journal of Education*, *2*(2), 49–57. https://doi.org/10.5430/wje.v2n2p49

Fathman, A. K., & Whalley, E. (2012). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second language writing* (pp. 178–190). Cambridge University Press.

Fazilatfar, A. M., Fallah, N., Hamavandi, M., & Rostamian, M. (2014). The effect of unfocused written corrective feedback on syntactic and lexical complexity of L2 writing. *Procedia - Social and Behavioral Sciences*, *98*, 482–488. https://doi.org/10.1016/j.sbspro.2014.03.443

Fazio, L. L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority- and majority-language students. *Journal of Second Language Writing*, *10*(4), 235–249. https://doi.org/10.1016/s1060-3743(01)00042-x

Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, *8*(1), 1–11. https://doi.org/10.1016/s1060-3743(99)80110-6

Ferris, D. (2003). *Response to student writing: Implications for second language students*. Lawrence Erlbaum Associates.

Ferris, D. (2006). Does error feedback help student writers? new evidence on the short-and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81–104). Cambridge University Press.

Ferris, D. (2012). Written corrective feedback in second language acquisition and writing studies. *Language Teaching*, *45*(4), 446–459. https://doi.org/10.1017/s0261444812000250

Ferris, D. R. (1995). Student reactions to teacher response in multiple-draft composition classrooms. *TESOL Quarterly*, *29*(1), 33–53. https://doi.org/10.2307/3587804

Ferris, D. R., & Hedgcock, J. S. (2005). *Teaching ESL composition: Purpose, process, and practice*. Lawrence Erlbaum Associates.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*(3), 299–323. https://doi:10.1017/S0272263100015047

Grimes, D., & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. Journal of Technology, *Learning, and Assessment*, *8*(6), 1–43.

Hajebi, M. (2018). Enhancing writing performance of Iranian EFL university students in the light of using computer-assisted language learning. *International Linguistics Research*, *1*(2), 47–51. https://doi.org/10.30560/ilr.v1n2p47

Huang, S., & Renandya, W. A. (2018). Exploring the integration of automated feedback among lower-proficiency EFL learners. *Innovation in Language Learning and Teaching*, *14*(1), 15–26. https://doi.org/10.1080/17501229.2018.1471083

Huxham, M. (2007). Fast and effective feedback: Are model answers the answer? *Assessment & Evaluation in Higher Education*, *32*(6), 601–611. https://doi.org/10.1080/02602930601116946

Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of Second Language Writing*, *7*(3), 255–286. https://doi.org/10.1016/s1060-3743(98)90017-0

Hyland, K. (1993). ESL computer writers: What can we do to help? *System*, *21*(1), 21–30. https://doi.org/10.1016/0346-251x(93)90004-z

Kaivanpanah, S., Alavi, S. M., & Meschi, R. (2020). L2 writers' processing of teacher vs. computer-generated feedback. *Journal of English Language Teaching and Learning*, *12*(26), 175–215. https://doi.org/10.22034/ELT.2020.11472

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *Modern Language Journal*, *99*(1), 1–18. https://doi.org/10.1111/modl.12189

Lai, Y. (2010). Which do students prefer to evaluate their essays: Peers or computer program. *British Journal of Educational Technology*, *41*(3), 432–454. https://doi.org/10.1111/j.1467-8535.2009.00959.x

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, *27*(4), 590–619. https://doi.org/10.1093/applin/aml029

Liao, H.-C. (2015). using automated writing evaluation to reduce grammar errors in writing. *ELT Journal*, *70*(3), 308–319. https://doi.org/10.1093/elt/ccv058

Lu, X. (2019). An empirical study on the artificial intelligence writing evaluation system in China CET. *Big Data*, *7*(2), 121–129. https://doi.org/10.1089/big.2018.0151

Milton, J. (2006). Resource-rich web-based feedback: Helping learners become independent writers. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 123–139). Cambridge University Press.

Mohammad Hosseinpur, R. (2015). The impact of teaching summarizing on EFL learners' microgenetic development of summary writing. *Teaching as a Second Language Quarterly*, *34*(2), 69-92.

Mohammad Hosseinpur, R., Bagheri Nevisi, R., & Bahrani, S. (2018). The impact of sequence map. answering questions, and 3-2-1 technique on EFL learners' summary writing ability. *Teaching as a Second Language Quarterly*, *37*(2), 125-167.

Nassaji, H., & Kartchava, E. (2021*). The Cambridge handbook of corrective feedback in second language learning and teaching*. Cambridge University Press.

O'Neill, R., & Russell, A. (2018). Stop! grammar time: University students' perceptions of the automated feedback program Ginger. *Australasian Journal of Educational Technology*, *35*(1), 42–56. https://doi.org/10.14742/ajet.3795

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, *24*(4), 492–518. https://doi.org/10.1093/applin/24.4.492.

Polio, C., Fleck, C., & Leder, N. (1998). "If I only had more time:" ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing*, *7*(1), 43–68. https://doi.org/10.1016/s1060-3743(98)90005-4

Qassemzadeh, A., & Soleimani, H. (2016). The impact of feedback provision by Ginger software and teachers on learning passive structures by Iranian EFL learners. *Theory and Practice in Language Studies*, *6*(9), 1884–1894. https://doi.org/10.17507/tpls.0609.23

Saadi, Z. K., & Saadat, M. (2015). EFL Learners' writing accuracy: Effects of direct and metalinguistic electronic feedback. *Theory and Practice in Language Studies*, *5*(10), 20–53. https://doi.org/10.17507/tpls.0510.11

Salavatizadeh, M., & Tahriri, A. (2020). The effect of blended online automated feedback and teacher feedback on EFL learners' essay writing ability and perception. *Teaching as a Second Language Quarterly*, *39*(3.2), 181–225. https://doi.org/10.22099/jtls.2021.38753.2899

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. TESOL *Quarterly*, *41*(2), 255–283. https://doi.org/10.1002/j.1545-7249.2007.tb00059.x

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, *46*(2), 327–369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

University of Cambridge. (2007). *Key English test: Handbook for teachers*. UCLES.

Valizadeh, M. (2020). The effect of comprehensive written corrective feedback on EFL learners' written syntactic accuracy. *Advances in Language and Literary Studies*, *11*(1), 17–21. https://doi.org/10.7575/aiac.alls.v.11n.1p.17

Valizadeh, M., & Soltanpour, F. (2021). Focused direct corrective feedback: Effects on the elementary English learners' written syntactic complexity. *Eurasian Journal of Applied Linguistics*, *7* (1), 132–150. https://doi.org/10.32601/ejal.911207

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing: Effectiveness of comprehensive CF. *Language Learning*, *62*(1), 1–41. https://doi.org/10.1111/j.1467-9922.2011.00674.x

Wang, Z., & Han, F. (2022). The effects of teacher feedback and automated feedback on cognitive and psychological aspects of foreign language writing: A mixed-methods research. *Frontiers in Psychology*, *13*, 1–12. https://doi.org/10.3389/fpsyg.2022.909802

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 Monologic oral production. *Applied Linguistics*, *24*(1), 1–27. https://doi.org/10.1093/applin/24.1.1

Zamel, V. (1985). Responding to student writing. *TESOL Quarterly*, *19*(1), 79–101. https://doi.org/10.2307/3586773

Appendices
Appendix A
KET Reading and Writing Sample Test

Part 1

Questions 1 – 6

For each question, choose the correct answer.

1

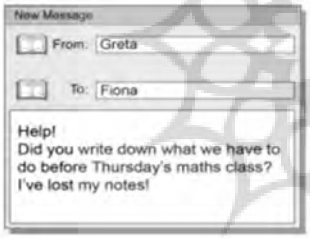
Go upstairs if you want to

A buy a dress for a party.

B pay less for something to read.

C find a game for a teenager.

2

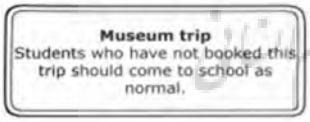


A Greta has forgotten when the next maths class is.

B Greta hopes Fiona will help her find her maths notes.

C Greta wants to know what the maths homework is.

3

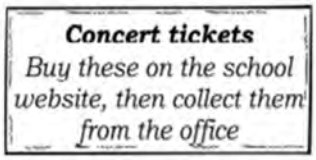


A Students not going on the trip cannot have a day off school.

B Students have to decide today if they would like to join the trip.

C Students going on the trip must come to school first.

4

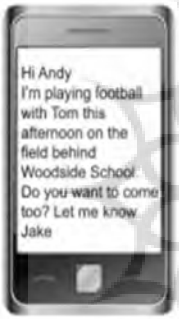


A Pay for tickets online before picking them up at school.

B Check the website for information about when tickets will be available.

C Let the office know soon if you are planning to buy tickets.

5



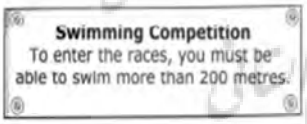
What should Andy do?

A invite some friends to play football

B tell Jake if he can join him later

C show Tom where Woodside School is

6



A Swimmers at all levels can enter this competition.

B This competition is for people who can swim over 200 metres.

C The races in the competition will be 200 metres long.

Part 2

Questions 7 – 13

For each question, choose the correct answer.

		Amy	Flora	Louisa
7	Whose class learnt about the garden competition from a TV programme?	A	B	C
8	Whose class grew some vegetables?	A	B	C
9	Whose class won a trip in the school garden competition?	A	B	C
10	Whose class painted flowers on their garden wall?	A	B	C
11	Whose class learnt about the insects in their garden?	A	B	C
12	Whose class got help from someone in a pupil's family?	A	B	C
13	Whose class chose flowers that were the same colour?	A	B	C

School gardens competition

Amy

Our class has just won a prize for our school garden in a competition – and they're going to make a TV film about it! The judges liked our garden because the flowers are all different colours – and we painted some more on the wall around it. My cousin gave us advice about what to grow – she's learning about gardening at college. We're planning to grow some vegetables next year. I just hope the insects don't eat them all!

Flora

Our teacher heard about the school garden competition on TV and told us about it. We decided to enter and won second prize! There's a high wall in our garden where many red and yellow climbing flowers grow and it looks as pretty as a painting! Our prize is a visit to a special garden where there are lots of butterflies and other insects. My aunt works there and she says it's amazing.

Louisa

The garden our class entered in the competition is very special. The flowers we've grown are all yellow! They look lovely on the video we made of the garden. We also grew lots of carrots and potatoes, and everyone says they taste fantastic. It was an interesting project. Our teacher taught us lots of things about the butterflies in our garden. We also watched a TV programme about them, and did some paintings to put on the classroom wall.

Part 3

Questions 14 – 18

For each question, choose the correct answer.

Starting at a new school

By Anna Gray, age 11

I've just finished my first week at a new school and I'd like to tell you about it. Like other children in my country, I went to primary school until I was eleven and then I had to go to a different school for older children. I loved my primary school but I was excited to move to a new school.

It was very strange on our first day. There were some kids from my primary school there, but most of the children in my year group were from different schools. But I soon started talking to the girl who was sitting beside me in maths. She lives near me so we walked home together. We're best friends now.

When I saw our timetable there were lots of subjects, some were quite new to me! Lessons are harder now. They're longer and the subjects are more difficult, but the teachers help us a lot.

At primary school we had all our lessons in one classroom. Now each subject is taught in a different room. It was difficult to find the classrooms at first because the school is so big. But the teachers gave us each a map of the school, so it's getting easier now.

The worst thing is that I have lots more homework to do now. Some of it is fun but I need to get better at remembering when I have to give different pieces of work to the teachers!

14 How did Anna feel about moving to a new school?

A worried about being with lots of older children

B happy about the idea of doing something different

C pleased because she was bored at her primary school

15 Who has become Anna's best friend at her new school?

A someone from her primary school

B someone she knew from her home area

C someone she met in her new class

16 What does Anna say about the timetable at her new school?

A It includes subjects she didn't do at primary school.

B She has shorter lessons than she had at her old school.

C It is quite difficult to understand.

17 Why couldn't Anna find her classrooms?

A She couldn't read a map.

B There was little time between lessons.

C The school building was very large.

18 What does Anna say about the homework she has now?

A She gets more help from some teachers than others.

B She thinks it is the hardest part of school life.

C She remembers everything she's told to do.

Part 4

Questions 19 – 24

For each question, choose the correct answer.

Wivenhoe hotel

Wivenhoe is a beautiful hotel in the countryside, with many rooms and an excellent restaurant. However, there is a big **(19)** ..... between Wivenhoe and other hotels. Firstly, Wivenhoe is part of a university, and secondly, its staff are all teenagers.

In fact, Wivenhoe is a hotel school for young people who are **(20)** ..... to get jobs in the hotel or restaurant **(21)** ..... . The students learn by helping staff in a real hotel, while their teachers **(22)** ..... them carefully. They do everything, from making beds and cleaning bathrooms to preparing menus and **(23)** ..... the telephone.

Some British people may think that a hotel run by students is a rather strange idea, but many visitors say that Wivenhoe is the best hotel they have ever **(24)** ..... at.

19	**A**	change	**B**	variety	**C**	difference
20	**A**	knowing	**B**	hoping	**C**	explaining
21	**A**	business	**B**	work	**C**	career
22	**A**	see	**B**	look	**C**	watch
23	**A**	calling	**B**	answering	**C**	speaking
24	**A**	entered	**B**	stayed	**C**	gone

Part 5

Questions 25 – 30

For each question, write the correct answer.
Write **one** word for each gap.

Example: 0 for

From: Anita
To: Sasha

Thank you **(0)** ..... your email. Living in Canada sounds really great! I'm glad that you like **(25)** ..... new house. What's the weather like? **(26)** ..... it very cold in Canada? Does it snow every day?

I heard that a **(27)** ..... of Canadians speak two languages – English and French. Are you having French lessons? Do you watch programmes **(28)** ..... TV in French too?

How about the students in your new school? Are **(29)** ..... friendly? And send some photos too – I would like to know more about them.

I've got **(30)** ..... go now, but I'll write again soon.

Part 6

Question 31

You are going shopping with your English friend Pat tomorrow.
Write an email to Pat.

Say:

- where you want to meet
- what time you want to meet
- what you want to buy.

Write **25 words** or more.

Write the email on your answer sheet.

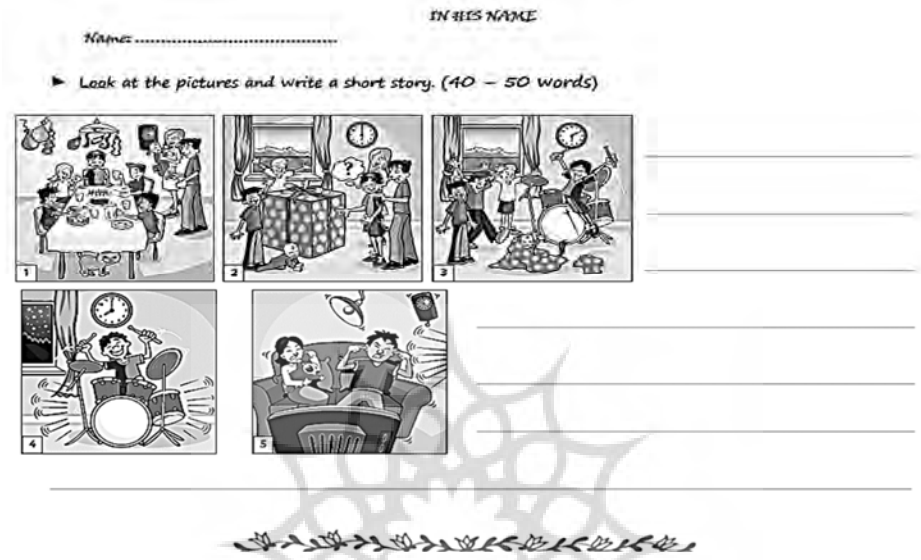
Part 7

Question 32

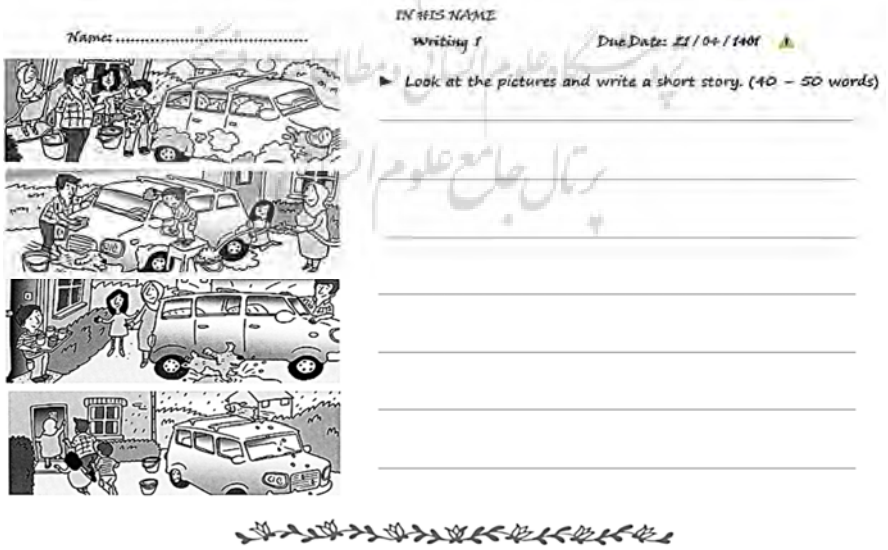
Look at the three pictures.
Write the story shown in the pictures.
Write **35 words** or more.

Write the story on your answer sheet.

Appendix B
Pre-test



Appendix C
1st Writing Homework



Appendix D
2nd Writing Homework

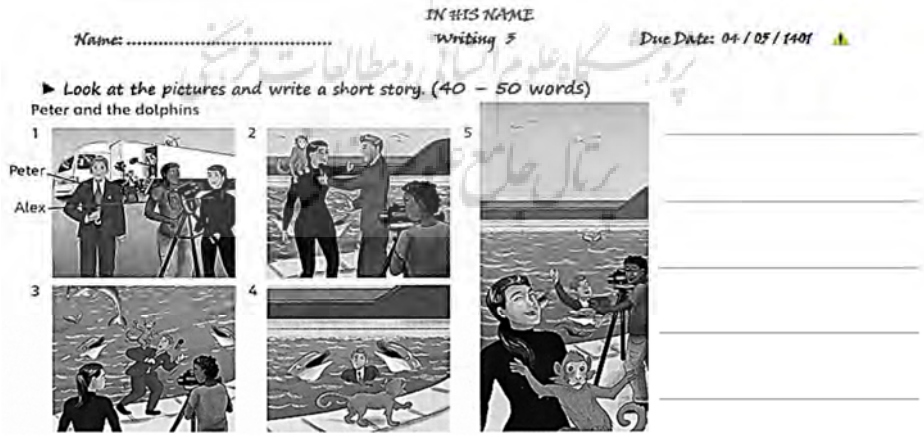

IN HIS NAME

Name: ..... Writing 2 Due Date: 25 / 04 / 1401

▶ Look at the pictures and write a short story. (40 – 50 words)

1 2 3 4 5

Appendix E
3rd Writing Homework


IN HIS NAME

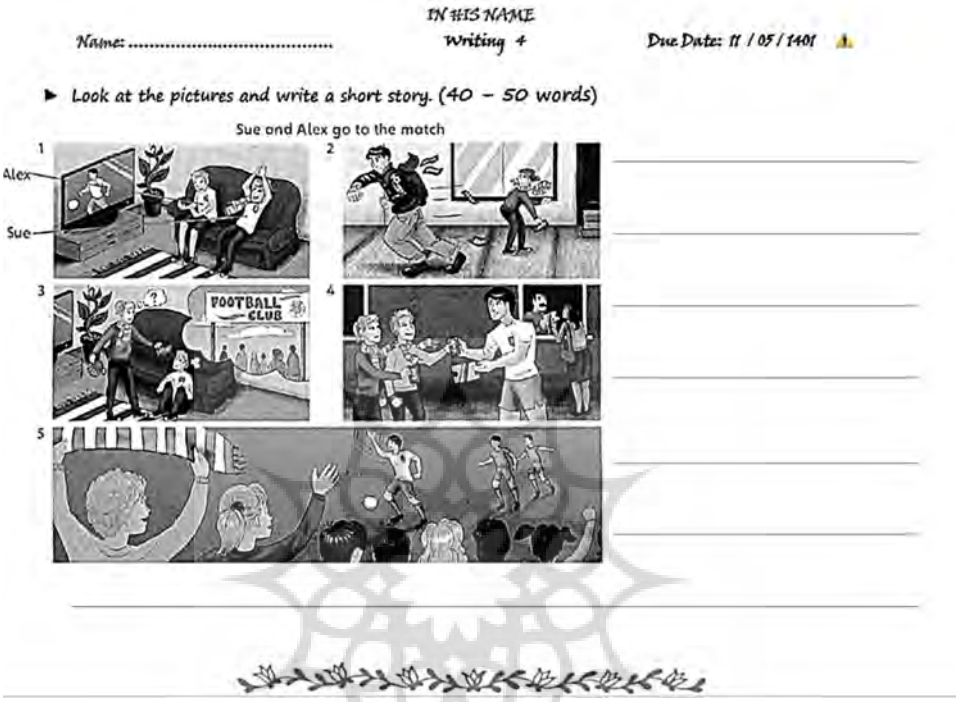
Name: ..... Writing 5 Due Date: 04 / 05 / 1401

▶ Look at the pictures and write a short story. (40 – 50 words)

Peter and the dolphins

1 Peter Alex 2 5 3 4

Appendix F
4th Writing Homework



Appendix G
Post-test

The Questionnaire of Group T

In the Name of God

My Idea about My teacher's Feedback and Suggestions

⁘ *Please make sure to answer all the questions. Your answers would be anonymous.*

		Strongly disagree	disagree	neutral	agree	Strongly agree
1	I feel satisfied with my teacher's feedback.	0	1	2	3	4
2	I feel satisfied with my teacher's comments on my writings.	0	1	2	3	4
3	I feel satisfied with my teacher's comments on my grammatical mistakes.	0	1	2	3	4
4	I feel satisfied with my teacher's comments on my vocabulary mistakes.	0	1	2	3	4
5	I feel satisfied with my teacher's comments on my spelling mistakes.	0	1	2	3	4
6	I do not feel satisfied with teacher's feedback on the sentences of my writings.	0	1	2	3	4
7	My teacher's feedback helps me reduce my mistakes in grammar.	0	1	2	3	4
8	My teacher's feedback helps me reduce my mistakes in Vocabulary.	0	1	2	3	4
9	My teacher's feedback helps me reduce my mistakes in spellings.	0	1	2	3	4
10	My teacher's feedback helps me reduce my writing mistakes.	0	1	2	3	4
11	My teacher's feedback does not help me improve my grammar.	0	1	2	3	4
12	My teacher's feedback does not help me learn some vocabularies.	0	1	2	3	4
13	My teacher's feedback does not help me improve my English writing.	0	1	2	3	4

		Strongly disagree	disagree	neutral	agree	Strongly agree
14	I can understand my teacher's feedback on my errors.	0	1	2	3	4
15	I can understand my teacher's suggestions.	0	1	2	3	4
16	I think my teacher's feedback is clear.	0	1	2	3	4
17	My teacher's suggestions are not understandable.	0	1	2	3	4
18	My teacher's handwriting is easy to read.	0	1	2	3	4
19	My teacher's feedback encouraged me to write more.	0	1	2	3	4
20	My teacher's feedback encouraged me to improve my writing.	0	1	2	3	4
21	My teacher's feedback does not encourage me to be a better writer.	0	1	2	3	4
22	My teacher's suggestions encouraged me to think about or write different sentences.	0	1	2	3	4
23	I like to receive my teacher's feedback again.	0	1	2	3	4
24	I do not want to receive my teacher's feedback again.	0	1	2	3	4
25	I do not want to receive my teacher's suggestions again.	0	1	2	3	4
26	I would like to receive my teacher's feedback and suggestions in the future.	0	1	2	3	4

Thank you!

Appendix I
The Questionnaire of Group C

In the Name of God

My idea about Ginger's Feedback

❖ *Please make sure to answer all the questions. Your answers would be anonymous.*

		Strongly disagree	disagree	neutral	agree	Strongly agree
1	I feel satisfied with Ginger's feedback.	0	1	2	3	4
2	I feel satisfied with Ginger's comments on my writings.	0	1	2	3	4
3	I feel satisfied with Ginger's comments on my grammatical mistakes.	0	1	2	3	4
4	I feel satisfied with Ginger's comments on my vocabulary mistakes.	0	1	2	3	4
5	I feel satisfied with Ginger's comments on my spelling mistakes.	0	1	2	3	4
6	I do not feel satisfied with Ginger's feedback on the sentences of my writings.	0	1	2	3	4
7	Writing with Ginger helps me reduce my mistakes in grammar.	0	1	2	3	4
8	Writing with Ginger helps me reduce my mistakes in Vocabulary.	0	1	2	3	4
9	Writing with Ginger helps me reduce my mistakes in spellings.	0	1	2	3	4
10	Writing with Ginger helps me reduce my writing mistakes.	0	1	2	3	4
11	Writing with Ginger does not help me improve my grammar.	0	1	2	3	4
12	Writing with Ginger does not help me learn some vocabularies.	0	1	2	3	4
13	Writing with Ginger does not help me improve my English writing.	0	1	2	3	4

		Strongly disagree	disagree	neutral	agree	Strongly agree
14	I can understand Ginger's feedback on my errors.	0	1	2	3	4
15	I can understand Ginger's suggestions.	0	1	2	3	4
16	I think Ginger's feedback is clear.	0	1	2	3	4
17	Ginger's suggestions are not understandable.	0	1	2	3	4
18	Working with Ginger encouraged me to write more.	0	1	2	3	4
19	Working with Ginger encouraged me to improve my writing.	0	1	2	3	4
20	Working with Ginger does not encourage me to be a better writer.	0	1	2	3	4
21	Working with Ginger Rephraser encouraged me to think about or write different sentences.	0	1	2	3	4
22	I will continue using Ginger.	0	1	2	3	4
23	I think it is good to receive Ginger's feedback before submit my writing to the teacher.	0	1	2	3	4
24	I will not work with Ginger again.	0	1	2	3	4
25	I will not work with Ginger Rephraser again.	0	1	2	3	4
26	I would like to use Ginger in the future.	0	1	2	3	4

Thank you!